International Journal of Economics & Strategic Management of Business Process (ESMB)
Vol.11 pp.1-14

Multilayer perceptron (MLP) and radial basis function network for the prediction of the Tunisian small and medium enterprises' (SME) bankruptcy.

Madiha Zammel[1]

Walid Khoufi[2]

Abstract

The objective of this work is to compare the development of neural statistical multilayer perception models and the radial basis function network for the prediction of the Tunisian small and medium enterprises' (SME) bankruptcy. The superiority of a multilayer perceptron (MLP) is easily detected. Our sample consists of 130 Tunisian companies and 18 financial ratios calculated for the 2005/2012 period. The forecasting results obtained from the neuronal model (MLP) are compared to the ones derived from the Linear Discriminant Analysis (LDA). The obtained results confirm the superiority of the neuronal technique in terms of predictability.

Keywords: corporate bankruptcy, forecasting, multilayer Perceptron, radial basis function, discriminant analysis.

1. Introduction

Corporate bankruptcy is a phenomenon that purifies the market of inefficient and unproductive firms. However, this reality is expensive for all the stakeholders such that the creditors may remain unpaid, employees may lose their jobs and owners will be at risk of losing their funds….. Therefore, acting before the outbreak of a problem certainly helps you avoid the risk of bankruptcy and save your costs. For this reason, forecasting is a major focus of the academic research studies which deal with bankruptcy not only theoretically but also practically.

Bankruptcy prediction, which consists in estimating the business default risk on the basis of the accounting data, has long been the major concern o researchers since the early 20th century, mainly with the work of Rosendale (1908), Fitz Patrick (1932).It consists in using statistical tools to detect the emergence of financial signals indicating a future failure. There is a variety of models that can be used to predict bankruptcy. Generally, the common objective of these models is to assign any business to either group (a group of bankrupt companies and another of healthy ones) using accounting ratios. Although the methodology used in these models and the analysis variables have not much changed, the evolution of the statistical analysis tools made this topic always relevant. In fact, the historical development of the prediction models may be divided into two major periods: the time of the traditional models

[1]PhDstudentattheFacultyof Economics and Management of Sfax: Governance, FinanceandAccountinglaboratory. E-mail: madiha-zammel@hotmail.fr

[2]Professor atthe Institute of Higher Commercial Studies of Sfax. E-mail: walid.koufi@yahoo.fr

ISSN 2356-5608

and that of the artificial intelligence ones. Moreover, traditional studies about bankruptcy prediction can be classified into two groups according to the existence or absence of hypotheses about the distribution of the variables. If the distribution law of the variables is known in advance, one therefore can speak about parametric statistical methods, whereas in the opposite case, one can speak about non parametric statistical methods.

First of all, the parametric statistical classification methods are intended to establish a functional relationship the form of which is given a priori between an endogenous variable and the exogenous variables the distribution law of which is assumed to be known. The parametric prediction models can be grouped into three families. At first, the authors used a single variable to explain the business failure, hence, one can speak about a univariate statistical analysis (Fitz Patrick (1932), Winakor and Smith (1935), Merwin (19 42), Beaver (1966)). Then, since the variables multiply with the hypothesis and follow the normal distribution, one can therefore speak about a multi-varied discriminant analysis (Altman (1968)). Finally, the hypothesis of multi-varied normality of the variables is rejected in favor of the hypothesis which states that the explanatory variables have different distributions, hence the birth of the logistic regression methods (Probit/Logit) (Ohlson (1980)).

Actually, when the financial ratios do not follow a multi-normal distribution and the errors follow neither the normal nor the logistic distribution, predicting bankruptcy using parametric models is impossible, hence the resort to non-parametric models, which requires no hypotheses about the distribution of the variables. Then nonparametric techniques can be classified into two categories: models based on recursive partitioning technique (Frydman et al (1985), Bardos (1989)) and models based on the kernel estimator technique (Calia and Ganugi (1997)).

The second stage of history began with the algorithms of artificial intelligence which were developed thanks to the many constraints of traditional statistical methods. The method of artificial neural networks is one os the most famous techniques in this research field. This nonparametric technique has its origins in the artificial intelligence, specifically, in the sector on machine learning (Remade. C 2004). The techniques of artificial intelligence have produced significant results about bankruptcy forecasting without requiring any statistical restriction. The several studies that focused on the comparison between these new techniques and the traditional ones confirmed their superiority regarding the forecasting quality (Mamoghli and Jallouli (2006)). Moreover, the artificial intelligence techniques have the advantage of being adapted to any raised problem, including corporate bankruptcy prediction.

ISSN 2356-5608

Taking the Tunisian context, where corporate bankruptcy has been on the rise, we suggest applying the multilayer perceptron model as well as the radial basis model for the prediction of corporate bankruptcy for a sample of 130 companies. For this purpose, the second section of our contribution will be devoted to the presentation of MLP and RBF models: the network architecture, the transfer function and the learning algorithm. The third section is confined to the application of these models to the bankruptcy prediction of the Tunisian companies. This section concludes with a comparative study between the performance of the multilayer perceptron model and that of the discriminant analysis.

2. Models of artificial neural networks: MLP and EBF

An artificial neural network (ARN) is a non parametric calculation tool inspired by the neural biological system made up of simple mathematical operators called "formal neurons". In fact, the formal neuron is a non-linear algebraic and *set* up function of the real variables. The development of the technique of the artificial neurons comes from an imitation of some mechanisms of the human brain. A neural network is a set of interconnected units, which have a large learning and processing information capacity. In fact, it is, a mathematical algorithm that can perfectly handle the knowledge related to the relationship between the input and output values so as to correctly classify the situations. The neural network models are quite diversified, which makes the criteria of the model classification multifold. These classification criteria include: the network architecture, the (supervised or unsupervised) learning algorithm, the combination function.... In the models of the neural networks, each neuron of the hidden and output layers receives neuron values beforehand through these synaptic connections and thus produces its own value by using its combination function. Based on this function, one can single out two types of networks: MLP networks (multilayer perceptron) and radial basis function networks (RBF).

2.1. The MLP model

The multilayer perceptron model (MLP) is a static structure of the neural network which does not present feedback loops but the learning of which is supervised.

2.1.1. Network architecture

The multilayer perceptron is a network composed of 3 types of successive layers. Firstly, an input layer which brings together incoming signals where each neuron is used as an input. Secondly, there is a set of hidden and positioned layers to participate in the transfer. It is worth noting at this level that each hidden layer neuron is connected at the entry to each neuron of the previous layer and at the exit to

ISSN 2356-5608

each neuron of the ollowing layer. The MLP is a network of "feed-forward" type where one layer can use the outputs of the previous layers (acyclic). In general, the MLP can have any number of layers and neurons per layer. The neurons are interconnected by weighted connections. In fact, the role of these connections is to use their weight to program an application of the entrance space to the exit space using a non-linear transformation.

2.1.2. The transfer function

The MLP network achieves a scalar product between its input vector 'X' and the 'W' weight vector, adds a 'b' bias and uses an activation function 'F' to determine its output. Hence, the neuron 'i' receiving information from neurons in the previous layer, performs the following operation:

$n_i = F(S_i)$ et $S_i = \sum_{i=1}^{n} w_{ij} n_j$

with :

n_i : is the neuron state

n_j : is the state of neuron j which preceded i

w_{ij} : the synaptic weights

F : the transfer function

Different transfer functions, such as the linear function, the standard sigmoid function and mainly the hyperbolic tangent function, can be used in this type of network,

2.1.1. Learning

The learning principle in the MLP models is to know the contribution of each weight in the global network error. Actually, there are several methods to perform learning among which we can mention the conjugate gradient algorithm, the methods of the second order and the spread of the retro gradient algorithm which is most used in the academic research studies.In fact, the gradient of the propagation algorithm measures the error between the desired outputs and the observed outputs that result from the spread forward entries and retro propagates this error through the network layers by moving from the outputs towards the inputs. Hence, the weight isadjusted for the neurons one by one starting with the output layer.It is worth noting that at the MLP level, the rise of the number of hidden layers only adds the progress of the learning algorithm that requires more iterations to converge to a result. For this reason, the number of hidden layers in an MLP does not usually exceed two. The learning algorithm ends and then the network converges when the weights are not changed or their modification is very low.

2.2. The radial basis function model (RBF)

The radial basis neural networks (RBF) are a special class of multi-layer neural networks (F. Yang and M. Paindavoine 2005). This is a

ISSN 2356-5608

powerful Feed-forward architecture with a supervised learning algorithm. RBF networks are mostly used to solve problems in large spaces.

2.2.1. Network architecture

The RBF network consists of three layers; a linear activation function input layer that transmits the entries without modification, a single hidden layer containing RBF neurons which are usually Gaussian functions, and a layer of output neurons that are typically activated by a linear activation function. Each layer is completely connected to the next one but there is no connection between the neurons of a same layer. Similar to the MLP, the RBF model can be used in the prediction, the classification....

2.2.2. The transfer function: radial basis function

The radial basis functions emerged at the end of the 1980s as new forms of neural networks. Various types of functions can be used as basic functions. However, the Gaussian function remains the most used. Hence, the general form of this function is as follows:

$$H(y) = \text{Exp} (- |y - c|^2 \ / \ 2\ \delta^2)$$

with :

C : is the center of the Gaussian function

δ: the dispersion

2.2.3. Learning

The learning process of the radial basis functions starts with the calculation of the Gaussian centers using different techniques, such as the K-means method or simply the calculation of the average arithmetic mean of the input vectors. Instead of achieving a weighted sum of these inputs, this type of network compares the inputs to the computed Gaussian centers. Each Xi input is therefore associated with a Ci value. The comparison of these two quantities is usually made using the Euclidean distance. In other words, the neuron starts calculating the following size:

$$d = \|X - C\| = \sqrt{\sum_{i=1}^{n} (Xi - Ci)^2}$$

Then, the neuron transforms the value achieved by a non-linear activation function of the Gaussian-type:

$$H (d) = \text{Exp} (- d^2 \ / \ 2\ \delta^2) = Z$$

With δ is the standard deviation of the activation function. An empirical rule is to compute δ using the following function:

$$\delta = \left(\frac{dn}{\sqrt{2M}} \right)$$

With:

d_n = the maximum distance between the n center and the other centers;

M = the total number of centers of the hidden layer

ISSN 2356-5608

Finally, the network output reads as follows:

$$S = f\left(\sum_{i=1}^{n} W_j Z_j + b_j\right)$$

With:

f : a linear function;

WJ: the synaptic weight between the neurons in the hidden layer and the output layer;

Zj: the output of the hidden layer

bj: the network bias.

2.3. Theoretical comparison between both MLP and RBF models

An MLP and RBF are two models of artificial neural networks which are very similar. Due to this similarity, it can be said that the RBF is a special case of an MLP for a hidden layer. The similar points between these two models are multiple. First, the output function, in most cases, is a simple linear function (except for a few specific problems, it may be another function) which returns a weighted sum of the values calculated through the neurons of the hidden layer. Then, the connections between the layers have the same meaning where each neuron is fully connected to the next layer neurons. Finally, the learning algorithm used for both models is supervised.Despite these broad similarities between the RBF and MLP model, some remarkable differences should be mentioned. First, the number of hidden layers is unique for an RBF for any problem to study but can be multiple for an MLP. Second, an RBF model uses only the radial basis function whereas an MLP may use any other function as an activation function.

3. Implementation of the MLP and RBF models to the failure prediction of the Tunisian companies

In what follows, we apply multilayer perceptron neuron networks and the radial basis functions networks to predict the failure of the Tunisian SMEs. For the MLP model, we will test the different architectures to finally choose the best and compare it to an RBF model. Finally, we carry out a comparison of the performance relative to the best-performing model (MLP or RBF) with that of the discriminant analysis.

3.1. The working methodology

3.1.1. The database and the variables

In this study, we prepared prepare our own database which includes 130 small and medium healthy Tunisian companies 65 of which are failing and 65 are healthy. During the construction of the sample, we met some homogeneity criteria. First, this is about the non-financial companies where only firms belonging to the industrial and commercial sectors are retained, which ensures the sector homogeneity.

ISSN 2356-5608

Then, the firms that make up our sample are small and medium-sized (SMEs), which brings back us to speak of the size uniformity. To define an SME, and according to the availability of data, we have chosen the size criterion according to some earlier studies, Kapuria and Faulkner (2008). Actually, an SME is a company the capital of which does not exceed 1 million Tunisian dinars. Finally, the homogeneity of groups of companies is met through the use of the paired sampling technique which consists in associating a healthy company with each defaulting one having the same size and sector of activity.

The sample of healthy enterprises is composed of companies that have no financial problems. However, for the construction of the sub-sample of the failing companies, and due to the multiplicity of definitions of the notion of failure, we have chosen to consider failing any company that enters into legal proceedings. The financial statements (balance sheets and income statements) of these companies are collected from the services of the Central Bank, the INS and control offices. The collected financial statements are about the 3 years preceding the failure year for the period between 2005 and 2012. In this way, we obtained a database over three- year period about 130 companies.

To analyze the financial situation of the firms in our sample, we set up a range of 32 financial ratios. The descriptive analysis of these ratios as well as the calculation of the correlation matrix enabled us to eliminate ratios that are correlated in absolute value greater than 0.7. Finally, we got 18 correlated ratios classified into 4 categories: liquidity ratios, management ratios, profitability ratios, and structure ratios (appendix 1). These ratios will serve as inputs for our Neuron network models.

3.1.2. Data configuration

The explanatory variables (financial ratios) data are unprocessed gross values having different size. To standardize the measurement scales, these data are converted into standardized values. Indeed, the Xi values of each ratio are standardized compared to their average and their standard deviation using the following relationship:

$$X_i' = \frac{Xi - \mu i}{\delta i}$$

With

XI' = the standardized value relative to R_i ratio

Xi = the gross value relative to R_i ratio

μi = the average value relative to R_i ratio

δi = the standard deviation relative to R_i ratio

ISSN 2356-5608

3.2. Results and discussions

3.2.1. Development of the MLP model

In what follows, we will apply the MLP multilayer networks. This choice can be explained by the ability of this model to meet the specificities of the prediction problem. Rizwan et al. (2014) think that this network can achieve very significant results and has a high prediction accuracy. An MLP model consists of a collection of neurons forming n-layer. The first layer is the vector of the input data whereas the last one contains the output vector and in between a number of hidden layers are positioned. The number of hidden layers determines the degree of network complexity. To identify the optimal network architecture to be used, we first chose the backpropagation algorithm of the gradient as the algorithm learning due to its frequency use in the prediction field (Mamoghli and Adele (2006), WU et al..)(2007), Bose et Pal (2006), Yim et Mitchell (2005), Tang et Chi (2005), West et al. (2005), Min et Lee (2005), Shin et al. (2005), Lee (2004), Kim et Han (2003), Agarwal et al. (2001), Atia (2001), Tan et Dihardjo (2001)…) . Then, for the transfer function at the level of the hidden layers, we have chosen the hyperbolic tangent function which is considered to be ideal for the MLP since it gives better results due to its symmetry:

Hyperbolic tangent function= $\frac{Exp\ (x)-Exp\,(-x)}{Exp\,(x)+\,Exp\ (-x)}$

Finally, the average of the squared error is used as a performance function:

$$MSE\ = ½\ \sum_{i=1}^{n}\quad(\ Vd_i - Vc_i\ )^2$$

With :

n : the number of observations

Vd_i : the output appreciated value

Vc_i : the output calculated value

However, we have chosen to test two functions of activation for the output layer, such as the identity function and the hyperbolic tangent function. Regarding the hidden layers and given that several authors (Altman et al. 1994, Desmet 1996) claimed that a neural network requires no more than two hidden layers, we have chosen to practice the learning base on a network with a single hidden layer and then on a network with two hidden layers. Regarding the number of neurons at the level of the hidden layers, there is no exact rule that limits it. However, Yao et al. (1999) proposed an iterative method which consists in gradually increasing the number of hidden neurons of a unit until (n 1) is reached, where n is the number of neurons in the previous layer. The 20 SPSS software is our computer tool for the implementation of this work. This software includes a 'Neuron network' command which helped us model the neural networks, and a subcommand «Perceptron multistarte» for the modelling of the multilayer perceptrons.

ISSN 2356-5608

3.2.2. Results of the MLP model

To determine the best network architecture, we varied the number of the hidden layers as well as the activation function for the output layer. For this reason, we randomly divided our database into two parts; 70% for the learning sample and 30% for the test sample. The following table summarizes the different tested architectures

Table1: The tested architectures

	Architecture : A1	Architecture : A2	Architecture : A3	Architecture : A4
Nombre des CC*	1	1	2	2
Transfer Fonction for CC	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent
Transfer Fonction for CS**	Hyperbolic tangent	Identity	Hyperbolic tangent	Identity

*CC : hidden layer **CS : output layer

The results of analysis of these 4 architectures using the SPSS are summarized in the following table:

Table2: Results of the test of an MLP architectures

	A1	A2	A3	A4
MSE	22,12	24,85	37,48	22,83
TBC**	93,8 %	88,3 %	79,7 %	90,4 %

**TBC : Rate of a good ranking

From these results, we conclude that the A1 architecture, which is a single hidden layer with the hyperbolic tangent function as an activation function for the hidden layer neurons and the output layer, is the best because it gives the highest TBC (93.8%), and the level of the lowest error (22.12). For this reason, in our case of failure prediction of the Tunisian companies using a multilayer perceptron model, the most efficient network is a three-layer network which uses the algorithm of back-propagation of the gradient as a learning algorithm, the hyperbolic tangent as an activation function for the output hidden layers, and the average quadratic error function as a performance function. This network configuration is [18-8-2], that is, 18 neurons in the input layer that represent our 18 financial ratios, 8 neurons in the hidden layer which establish the internal calculation of the network, 2 neurons of the output layer that represent the situation of the company either defaulting or healthy.

3.2.3. Development of the radial base model

The RBF neuron network consists of an input layer containing 18 neurons that represent our data entry, a hidden layer, and an output layer containing two neurons indicating the status of the failing or healthy company. The input variables consist of 18 independent and normalized vectors between -1 and 1 using the following relationship:

ISSN 2356-5608

Standardized X = [2* ($\frac{X - \min}{\max - \min}$)] - 1

These 18 input vectors represent 18 calculated financial ratios. For the development of the RBF network, we begin by calculating the Gaussian center of each C_1… C_{18} vector, using the mean arithmetic method as well as the Euclidean distance between the 18 vectors and their centers. Then, the Euclidean distance is transformed using a Gaussian function at the level of the hidden layer. The outputs of the neurons of the hidden layer will by the identity function. The development of an RBF model using the SPSS 20software helped us reach the architecture [18-6-2]. The results of this model are presented in the following table:

Table be weighted using the Wi weights, then added to the bias of the output layer, and finally transformed 3: Results of the RBF model

	RBF Model
Error	52,48
TBC	73,6 %

3.2.4. Comparison between MLP and RBF

The comparison between the results of the developed MLP and RBF models shows that the MLP represents the best performance. Moreover, it gives the highest good ranking rate (93.8% 73.6%). On the other hand, its learning phase is the fastest and converges for a reduced number of iterations compared to the RBF network. Several previous studies reached similar results that confirm the superiority of the MLP model compared to the RBF models. Thesestudiesinclude: Simani et al (2000), Mansouri et al (2008), Mrabti et al (2009) et Badaoui et al (2014).

3.2.5. Comparison between MLP and ADL

To improve the prediction quality, the linear discriminant analysis function is to globally assess the financial situation of the company described by a range of conjunction ratios. In fact, it is to determine a discriminant function for each company and decide a Z-score * to carry out the ranking. Actually, the classification rule is:

$$Z = \beta + \alpha_1 x_1 + \ldots.. + \alpha_n x_n$$

If Zi ≥ Z * then, the company is said to be healthy

ifZi Z * then, the company is considered to be faulting

The application of the method of the linear discriminant analysis (LDA) to our database first gives us the equality test results of the average groups that give information about the most discriminating variables. Therefore, these variables must take significant and high values in Fisher’s test.

ISSN 2356-5608

Table4: Equality test of the average groups

	Lambda de Wilks	F	ddl1	ddl2	Significance
R3	,976	4,841	1	196	,029
R10	,866	30,357	1	196	,000
R13	,949	10,539	1	196	,001
R16	,864	30,830	1	196	,000
R17	,947	10,983	1	196	,001
R21	,949	10,543	1	196	,001
R22	,899	21,952	1	196	,000
R24	,960	8,123	1	196	,005
R28	,899	22,055	1	196	,000

This table shows the relevance of the R16 ratios (coverage rate of the staff expenses), R10 (debt ratio), R22 (economic profitability ratio) and R28 (ratio of debt in the medium and long term). In fact, the debt ratio is a key factor in the explanation of failure. It affects the company's situation and helps discriminate between both groups of failing and healthy companies. In this context, Lelogeais (2003) thinks that usually failing companies suffer from a high level of short-term debt and low financial autonomy. In addition, the high cost of staff can worsen the financial situation of the company. Finally, the economic profitability ratio is generally a key element in measuring the company's economic performance and a generally significant indicator of its beneficiary capacity.

Our discriminant equation is as follows:

Z = 0,694 + (-0,749) R10 + 8,501 R16 + 0,338 R22 + (-3,220) R28

The classification result gives us a rate of correct classification equal to 82.3%.

The comparison of the multi-layer perceptron model and the discriminant analysis model shows the performance of the former fro on the basis the latter. In fact, the rate of the companies' good ranking rate provided by the MLP model is higher than that of the ADL model

Table5: The rate of good ranking of both models

TBC	
MLP	ADL
93,8 %	82,3 %

Consequently, based on the accurate ranking, the performance MLP is considered as a tool for predicting the model of business failure. It should be noted that this result is confirmed by several previous studies, such as Odom et Sharda (1990), Zhang et al, (1999), John et al (2000), Mamoghli et Jalouli (2006), Abdou et al, (2008)… .However, the superiority of the MLP model is not absolute since, according to some other authors, such as Boujelbene and Khmakhem (2007), these two models are complementary. Indeed, the internal links on a MLP have no economic meaning whereas the weights of the ratios in an ADL model are transparent and easy to interpret. Thus, complementarity lies in the ability of the LDA to

ISSN 2356-5608

select the most relevant variables and the use of an RNA template with these variables to calculate the lowest error rate.

3. Conclusion

Artificial neural networks are powerful tools of business failure prediction.Supervised learning, such as the MLP and RBF networks, showed a strong prediction ability. By comparing the results of these two models in this study, it can be concluded that the MLP model gives a higher TBC than that of the RBF. In fact, its learning phase is the fastest, besides, it converges for a limited number of iterations as the RBF model.Compared with the models of the artificial intelligence (NAS), the traditional methods, such as the ADL, remain applicable and provide convincing results. A comparison between an MLP and an ADL model helps us assess the superiority of the former based on the assessment of the ability of each model to correctly classify the surveyed companies. However, combining the error minimizing criterion and the ability of selection possibility of the most relevant variables, we could assess the complementarity both theMLP and ADL models.Finally, several other RNA types deserve to be considered to improve the quality of failure prediction. This is the model of non-supervised learning, such as the Kohonen map.

Appendix 1 : The used financial ratios

Code	Ratio
R1	*Currentassets/currentliabilities*
R3	*Cash and cash equivalent /current liabilities.*
R4	Permanent *capital / non-currentassets*
R8	Shareholders' equity/permanent capital
R10	*Total assets / total liabilities*
R12	*Net financial expenses / operating income*
R13	*Turnover / Total Assets*
R15	*Depreciationexpense/fixedassets*
R16	*Staff costs/added value*
R17	Customer and associated account* 365) / sales
R20	*Financial expenses* / *added value*
R21	*Net income* / net shareholders' equity
R22	*Operating income / total assets*

ISSN 2356-5608

R24	*Net fixedassets / total assets*
R25	*Total claims / total assets*
R28	*Medium and long term debt/ total assets*
R29	*Short term debt /total assets*
R32	*Log (total balance sheet.)*

Références

Abdou, H., Pointon, J. et El-Masry, A. (2008). Neural nets versus conventional techniques in credit scoring in Egyptian banking.*Expert Systems with Application, 35* (3), 1275-1292.

Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589-609.

Altman E.I., Marco G. etVaretto F. (1994). Corporate distress diagnosis: comparisons using linear discriminant and neural networks: the Italian Experience. *Journal of Banking and Finance, 20* (18), 505-529.

Badaoui, H., Abdallaoui, A. etChabaa, S. (2014).Perceptron Multicouches et réseau à Fonction de Base Radiale pour la prédiction du taux d'humidité.*International Journal of Innovation and ScientificResearch, 5* (1) , 55-67.

Bardos M. (1989). Trois méthodes d'analyse discriminante.*Cahiers Économiques et Monétaires, Banque de France,* 2ème trimestre, n°33.

Beaver W.H.(1966). Financial ratios as predictors of failure.*Journal of Accounting Research, supplément au vol. 4*, 71-111.

Calia, P. et Ganugi, P. (1997).Kernel and nearest neighbor discriminant analysis: Business failure classification in industrial district.AppliedStochasticModels and Data Analysis. Acte de Colloque Capri 1997.

Du Jardin, P. (2007). Prévision de la défaillance et réseaux de neurones : l'apport des méthodes numériques de sélection de variables.*thèse de doctorat*. Edition Dunod.

Rossi, F. (2005). Réseaux de neurones : le perceptron multi-couches. CoursUniversité Paris-IX Dauphine.

Fitz, P.(1932). A Comparison of Ratios of successful industrial enterprises with those of failed firms .Certified Public Accountant, October, November, and December, 1932, 598-605, 656-662, et 727-731.

Fouad, B. et Sylvie, T. (2003). Les Perceptrons Multicouches: de la régression non-linéaire aux problèmes inverses. Travail réalisée dans le cadre du programme NeuroSAT soutenu par la CEE.

Frydman, H., Altman E. et Kao D.L. (1985). Introducing recursive partitioning for financial classification: the case of financial distress.*Journal of Finance, 40* (1), 269-291.

Haykin, S. (1999).*Neural Networks-A Comprehensive Foundation*.Prentice-Hall, New York.

Lelogeais, L. (2003). Un score sur variables qualitatives pour la détection précoce des défaillances d'entreprises. *Bulletin de la Banque de France, n°114.*

ISSN 2356-5608

Mansouri, Y. Chetouani, B. et El Kihel.(2008). Using Neural Networks for fault detection in a distillation column.*International Journal Computer Applications in Technology, 32*(10), 181–186.

Merwin C. L. (1942). FinancingSmall Corporations in Five Manufacturing Industries.*National Bureau of Economic Research, Financial Research Program III, Studies in Business Financing.*

Mrabti, F. et Seridi, H. (2009).Comparaison de méthodes de classification réseau RBF, MLP et RVFLNN. *Damascus University Journal, 25*(2), 119-129.

Ohlson J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy.*Journal of AccountingResearch, 18*(1),109-131.

Refait C., (2004**).** La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux. *Economie et Prévision, 162*,129-147.

Rosendale W.M. (1908). Creditdepartmentmethods. *Banker's magazine,* 183-184.

Simani, S. et Fantuzzi, C. (2000). Fault diagnosis in power plant using neural networks.*Information Sciences, 127*(3), 125-136.

Koufi, W. et Feki, R.(2008). Détermination et analyse des caractéristiques des entreprises défaillantes à l'aide des cartes auto organisatrices de Kohonen (Cas des PMI Tunisiennes). La revue comptable et financière, recherche en comptabilité et finance, N°1.

Winakor A. H., et Smith R. F. (1935). Changes in the Financial Structure of Unsuccessful Industrial Corporations.University of Illinois, Bulletin n° 51.

Yang, F. et Paindavoine, M. (2005). Implantation d'un réseau de neurones RBF sur des systèmes embarqués : la détection et la reconnaissance de visages en temps réel.*RevueOptical Engineering, 44* (8),42-65 .

ISSN 2356-5608